International Journal of Core Engineering & Management

Volume-7, Issue-01, 2022 ISSN No: 2348-9510

COMPARATIVE ANALYSIS OF MACHINE LEARNING MODELS AND DATA ANALYTICS TECHNIQUES FOR FRAUD DETECTION IN BANKING SYSTEM

Siddharth Kumar Singh
New York University,
New York, NY USA
sks592@nyu.edu

Sanjeet Kumar Choudhary
Birla Institute of Technology,
Ranchi, India
sanjeet.choudhary1@gmail.com

Piyush Ranjan
Technology Arcitect
Cognizant , NJ USA
piyushranjangc@gmail.com

Sumit Dahiya
Apeejay College of Engineering, Gurgaon, India
sumitdahiya1234@gmail.com

Abstract

The expansion of the Internet has been phenomenal throughout the last ten years. The increased number of online payment options due to e-commerce and other websites has increased the risk of online fraud. Academics are starting to use a range of ML approaches for fraud detection and analysis in response to the increase in online transaction fraud. The main goal of this study is to use the Kaggle dataset to develop and deploy a novel method for detecting credit card fraud. ML models and data analytics strategies for banking fraud detection are compared in this article. It addresses the increasing sophistication of fraud schemes that financial institutions face and the need for advanced technologies to combat them. An evaluation is conducted to see how well ML models, such as XGBoost and Random Forest, can detect fraudulent transactions. Data preprocessing techniques, including normalization, feature selection, and class balancing, are employed to improve model performance. The findings show that both Random Forest and XGBoost outperform other models, achieving prediction accuracies of 99.95% and 99.96%, respectively, making them highly effective for fraud detection. In addition to strengthening the theoretical underpinnings of fraud detection, this study has important practical consequences for the financial industry, which stands to gain significantly from an improved fraud detection system.

Keywords: Machine Learning, Data Analytics Techniques, Fraud Detection, Credit card, Random Forest, XGBoost,

I. INTRODUCTION

The rapid expansion of the world's information infrastructure, especially in the domains of computing and IT (including telephone networks and the Internet), in the past few decades has catapulted electronic commerce to a global arena. The ability to communicate clearly and concisely with customers, competitors, and other organizations has been greatly enhanced by these innovations [1]. The goal of electronic commerce (e-commerce) is to facilitate the efficient transfer of data between companies, meet the demands of consumers, and gain a market advantage via the use of communication, data management, and security services. ICT is used in the banking industry to provide clients enhanced services and security, as it is in most other corporate domains[2]. Numerous customer-related services may be provided more easily due to their e-banking platform, which guarantees efficient contact between them and their clients. There are several names for e-banking in the literature, but they all relate to ICT-based financial transactions. These names include electronic banking, online banking, and virtual banking. Electronic banking, or e-banking, allows customers to access banking services from any location other than a physical bank branch [3][4].

Banking fraud detection is a serious issue as fraudulent activity is becoming a bigger hazard to financial organizations. There is a growing trend of CCF and other fraudulent acts in the financial industry[5]. Along with the exponential growth in the use of credit cards in everyday life, the fraud using these cards is also on the rise. Identity theft had its worst year ever in 2021, according to a study by the FTC, highlighting a gravity of a problem[6][7]. An important fact to keep in mind is that the stated numbers may not be reflective of the true prevalence of identity theft since many incidents go unreported. The FTC study highlights the need of finding new ways to protect organizations' and customers' financial security[8].

A rise of sophisticated fraud schemes has necessitated the use of advanced technologies to protect both banks and their customers[9]. An ability to analyses massive amounts of data in real-time using ML models and data analytics methods has made them formidable weapons in the fight against fraudulent transactions [10]. This paper explores various ML models and compares their performance in identifying fraud patterns. Additionally, it delves into data analytics techniques like anomaly detection, clustering, and predictive modeling, which are vital for detecting unusual transactions. Although there are a number of supervised ML methods available for fraud detection[11][12], our primary goal is to improve the dataset's capacity to handle a high volume of transactions, address the issue of significant class imbalance, and incorporate both labelled and unlabeled samples.

1.1 Motivation and Contribution of study

The increasing amount of digital transactions and the sophistication of fraudulent operations have led to a growing demand for effective fraud detection systems in the financial sector, which is the driving force behind our effort. Conventional techniques frequently fail to detect subtle patterns of fraud, particularly in datasets with extreme imbalances. This study attempts to improve fraud detection efficiency and accuracy by comparing sophisticated ML models like RF and XGBoost with data preparation methods like SMOTE and PCA. The goal is to improve the safety of the financial system by shedding light on the most effective models and techniques for detecting fraudulent transactions. In particular, the study's findings on applying ML methods to datasets with large imbalances advance our understanding of CCFD. Below are a contribution of the study:

- This article assesses the efficacy of several methods for dealing with data imbalance using the European credit card dataset on Kaggle.
- To tackle the problem of minority class fraud detection, it uses approaches such as under sampling and oversampling (SMOTE).
- Evaluates XGBoost and Random Forest, showcasing how well they detect fraudulent transactions.
- Uses confusion matrix and accuracy metrics to evaluate an effectiveness of fraud detection models.

1.2 Structure of the paper

Here is the structure of the remaining portion of the document: Section II delves into the relevant literature and its shortcomings, while Section III presents the technique and fraud detection system that has been suggested. Section IV outlines the experimental setting, analyses the data, and discusses the findings. Section V presents the conclusion and future suggestions.

II. LITERATURE REVIEW

This section provides the previous work on the banking fraud detection using ML. Fraud detection has received much attention in the past decade. Scholarly publications have observed the emergence of several research papers that have reviewed current techniques for FD and prevention.

Okuneye and Taiwo et al. (2018), assesses the many obstacles to preventing and identifying fraud in Nigeria's banking industry. According to the descriptive research, the most common kind of bank fraud in Nigeria is the theft of funds by bank directors and managers, rather than an inadequate level of motivation. In addition, the government should strengthen the anti-corruption agencies that are already in place and give them greater autonomy financially. As a warning to would-be con artists, managers and directors found guilty of embezzlement should face criminal charges [13].

Yee et al. (2018), explores the use of Bayesian network classifiers, specifically K2, TAN, Naïve Bayes, logistics, and J48 classifiers, for supervised analysis. In comparison to findings obtained before to dataset preparation, all classifiers produced results with an accuracy greater than 95.0% when normalization and PCA were applied to the dataset [14].

Thennakoon et al. (2019), primarily addresses four primary instances of fraud in actual financial transactions. Multiple ML models are applied to each fraud case, and the most effective strategy is chosen after review. They also evaluate an innovative approach that successfully deals with the data skew. Our experiments' data is sourced from a financial institution in accordance with a non-disclosure agreement [15].

Josephine Isabella et al. (2020), uses ML methods such SVM, NB, KNN, RF, DT, OneR, and AdaBoost to efficiently research FDS for credit cards. These ML methods assess a dataset and provide performance measures to determine their relative accuracy. The RFC was shown to be the most effective method out of all of the ones tested in this research [16].

Dave and Adewale et al. (2021), considers the potential benefits to Azerbaijan's banking industry of integrating behavioral analytics into fraud detection systems, with the aim of improving both security and efficiency. Within the context of Azerbaijan's legislative and technical environment, this study investigates the present fraud detection landscape, discusses the advantages of behavioral analytics, and outlines practical considerations for applying these techniques [17].

Jain et al. (2021), Luhn's algorithm and k-means clustering are employed to create a Credit Card Fraud Detection (CFD) system. In addition, FCM clustering, rather than k-means clustering, is also used in the development of CFD systems. They evaluate the two clustering methods' CFD performance using f-measure, recall, and precision. Results using k-means clustering are inferior to those from the FCM. To demonstrate the CFD system's performance even when faced with biassed data, additional assessment metrics are computed, including the rate of fraud detection, the rate of false alarms, the rate of balanced classification, and the Mathews statistical correlation coefficient[18]. The below table 1 provide the literature review summary with key way for fraud detection.

Table 1: Summary of the related work for fraud detection

Reference	Methodology	Results	Limitations	Future Work
[13]	Descriptive analysis of fraud causes in Nigerian banks	Lack of motivation is not a major fraud cause; looting by managers/directors is prevalent	Focuses only on Nigerian context; limited scope of analysis	Strengthening anti-graft agencies; exploring additional fraud detection methods
[17]	Behavioral analytics implementation in fraud detection systems	Enhanced security and efficiency through user behavior analysis	Regulatory and technological challenges in Azerbaijan	Further studies on integration of behavioral analytics with existing systems
[16]	Machine learning techniques (SVM, NB, KNN, RF, etc.) for CCFD	Random forest classifier outperformed others	Dataset limitations; may not generalize to all fraud types	Explore additional ML techniques and larger datasets
[15]	Evaluation of ML models for real-time CCFD	Comprehensive guide for selecting algorithms; predictive analytics used	Data skewness; reliance on confidential data	Develop strategies for handling skewed data more effectively
[14]	Supervised classification using Bayesian network classifiers	Achieved >95% accuracy post preprocessing	Dependence on specific preprocessing methods; may not generalize	Investigate other preprocessing techniques for accuracy improvement
[18]	CFD system using Luhn's algorithm and clustering techniques (k-means, FCM)	FCM clustering outperformed k-means in fraud detection metrics	Limited evaluation metrics; focus on credit card transactions only	Expand evaluation to include various types of fraud and other clustering algorithms

III. METHODOLOGY

The proliferation of ML techniques has coincided with an uptick in the creation of fraud detection systems powered by AI. The research relies on a number of ML models and data pretreatment strategies for identifying financial institution fraud. The dataset used for this research was sourced from Kaggle and included 284,807 credit card transactions, with a fraud rate of just 0.172%. A number of methods were used to rectify the severe class imbalance, including under sampling and the SMOTE. To ensure that the models concentrated on the most important features for fraud detection, preprocessing methods included data normalization and feature selection using PCA. Figure 1 depicts the whole procedure as a suggested flowchart for detecting fraud in the banking industry using ML models.

The flowchart each step discussed below briefly:

3.1 Data Gathering and data analysis

Credit card fraud statistics compiled from a European credit card firm. In order to distinguish among fraudulent and legitimate transactions, the dataset was divided into subsets according to the class label. The Kaggle platform is used to get the dataset. This dataset contains all of the credit card purchases made by cardholders in September of 2013. All purchases made during the last two days are included in the dataset. Out of the 284,807 transactions included in the dataset, 492 have been determined to be fraudulent. Only 0.172 percent of all transactions are fraudulent.

3.2 Data preprocessing

Data preparation is a crucial part of data mining. Preparing data for analysis involves processes such as cleaning, converting, and integrating. Improving the data's quality and making it more suited for the particular data mining activity is the purpose of data preparation.

- Handling missing value: Machine learning often has missing values. A variable without data points provides partial information and may compromise model correctness and reliability. Machine-learning programs must efficiently address missing values to yield powerful and impartial findings. Learn how to handle missing values in machine learning datasets in this article.
- Remove outlier: Outlier elimination is a popular preprocessing method. Noises are abnormal data points. Their presence in the training model may reduce classifier performance. These include removing irrelevant gestures too large or too small data.

3.3 Feature selection

Feature selection is a method for selecting useful, consistent, and non-redundant characteristics for use in building models. As datasets get more large and diverse, it becomes increasingly vital to reduce their sizes in a systematic manner. Feature selection's main goal is to boost a predictive model's performance while reducing the computational cost of modelling. Two crucial data variables, the amount of the transaction and the time, are shown by the distributions in Figure 2. There seems to be a significant rightward bias in the total amount of the transaction. You may see how many seconds have passed since the dataset's initial transaction with the help of the 'Time' feature. Exploratory data analysis included determining the extent to which the target categories were unequally represented in the data, in addition to searching for outliers and other anomalies. All three steps contribute to a better understanding of the data attributes, which in turn informs the data-processing operations required for data preparation, the next stage before modelling.

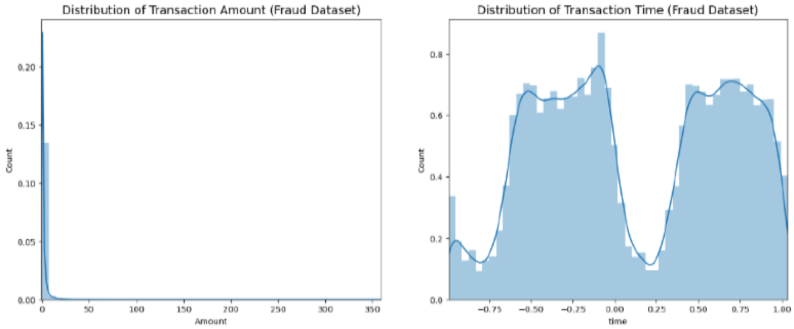

Figure 2: Distribution of transaction amounts and time.

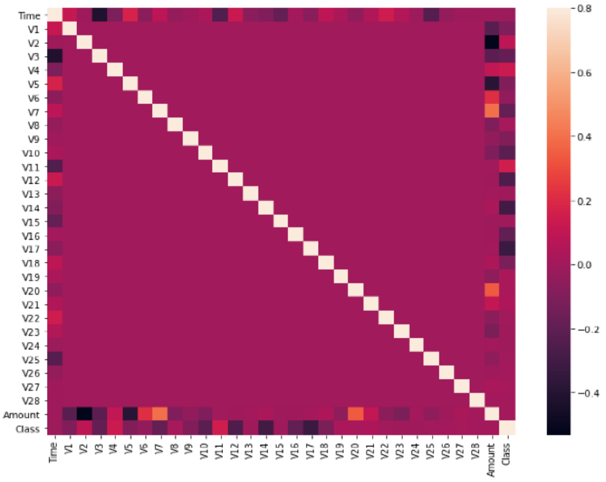

Figure 3: Heat map of Correlation matrix from fraud dataset

The heat map and correlation matrix are shown in Figure 3. An effective method that assists us in determining if a certain characteristic has to be removed is the correlation matrix. We conclude that there is no need to extract any features, and therefore, no need to preprocess the dataset, since the correlation matrix indicates that all characteristics, regardless of correlation strength, are connected to the 'Class' feature. Another explanation is that a PCA dimensionality reduction transformation produced the features from "V1" to "V28." Because sensitive information was included in the raw data for these characteristics, this action was taken.

3.4 Data Balancing with (SMOTE)

Some of the most popular oversampling techniques for dealing with imbalances are SMOTE (synthetic minority oversampling technique). Replicating members of minority groups at random increases their numbers, with the goal of achieving class parity. New minority instances are created by combining existing minority instances using SMOTE.

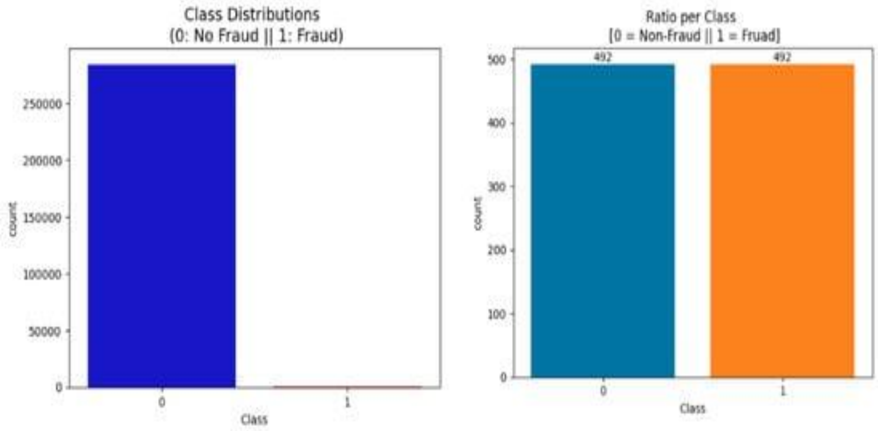

Figure 4: Target class distribution: (before) (imbalanced), (after) (balanced).

The dataset that was initially skewed and the one that was balanced following random under-sampling are both shown in Figure 4. One popular approach to dealing with big amounts of imbalanced data is random under-sampling. Reducing the size of the majority class by randomly removing samples is what it entails. The original statistics clearly show a significant class disparity, with fewer than 1% of transactions being fraudulent. Data variables have been filtered and the most relevant predictors have been chosen by applying dimensionality reduction, which is done using PCA.

3.5 Machine learning models

The most important aspect of our predictive modelling workflow is the last step, which involves using ML techniques to train and assess fraud prediction models. The following ML models are employed to categorize the credit card dataset's fraud detection performance.

1. **Random Forest**

The RF algorithm is an effective collective learning tool for regression and classification tasks. At the end of training, it produces the mean prediction for regression or the median of the classes for classification based on its many decision trees. To prevent overfitting and increase variety, the RF trains every decision tree using a different sample of the data and takes into account a different group of features at each split. The unpredictability and aggregation of numerous trees make random forests resistant to noisy data. Random forests prioritize feature relevance and resist overfitting, especially in high-dimensional data, making them effective for discovering significant predictors. Random forests are useful in fraud detection, healthcare, and finance due to their adaptability and usability.

2. **XGBOOST:**

XGBoost, is a scalable and quick gradient boosting method for ML experiments. By specializing on erroneously categorized or under-forecast occurrences, it builds an ensemble of decision trees that self-correct. XGBoost maximizes efficiency and speed via sparse data management, parallel processing, and regularization to avoid overfitting. It allows tree pruning, which stops branch growth when improvements are no longer possible, lowering model complexity. XGBoost's capacity to handle big datasets, missing values, and skewed data makes it popular in data science contests and real-world applications including fraud detection, recommendation systems, and risk assessment. Its versatility lets you tune hyper parameters to enhance performance, making it one of the most popular and capable machine learning models.

3.6 Performance matrix

For the model evaluation, use performance matrix like confusion matrix, and accuracy. The confusion matrix is a tool for evaluating the efficacy of ML classification on simulated data. A real class and a predicated class are its two fundamental components. For fraud detection in Banking System use accuracy for measures for determine the model efficiency.

Accuracy: We typically mean it when we claim something is correct, according to classification. A comparison is made between the amount of input samples and the percentage of right predictions. The following formula of accuracy is (equ.1):

$$Accuracy = \frac{TP + TN}{(TP + FN + TN + FP)} \dots \dots \dots \dots (1)$$

Precision: is the proportion of TP to the sum of all positive predictions (TP + FP (equ.2)) generated by a model. Simply said, it's the degree to which the model's optimistic predictions come true.

$$\text{Precision } = \frac{TP}{(TP + FP)} \dots \dots \dots \dots (2)$$

Recall: measures the accuracy with which a ML model locates all relevant occurrences of the positive class. This metric, which can be determined using equation 3, measures the proportion of positive observations that were correctly anticipated relative to the total numberof positive observations:

$$\text{Recall } = \frac{TP}{(TP + FN)} \dots \dots \dots \dots (3)$$

F1-score: is a statistic that takes the outcomes of recall and precision and puts them together into one number. The following equation (4) provides the formula for the F1-score:

$$\text{F1} - \text{Score} = 2 \times = \frac{Precision \times Recall}{Precision + Recall} \dots \dots \dots \dots (4)$$

An ambiguity matrix depicting the following is used to show the difference between trained and testing data:

- TP: True Positives are examples of training data when consumers were really a target of fraud and the results were accurate predictions.
- TN: The data that was not anticipated and does not correspond to the data that was manipulated is referred to as a true negative.
- FP: The data cannot be vulnerable to fraud, although a false positive is anticipated.
- FN: It is not possible to predict a false negative, but there is a real chance that the data is fraudulent.

IV. RESULT ANALYSIS AND DISCUSSION

The results of the ML models for finding scams should be analysed and talked about in this section. Computers with 64 GB of RAM and an Intel i7 core 158th generation are used to run the ML methods in the computer language Python [19]. Assess the effectiveness of ML models using the confusion matrix and measures of accuracy. The results of Two machine learning algorithms are as follows:

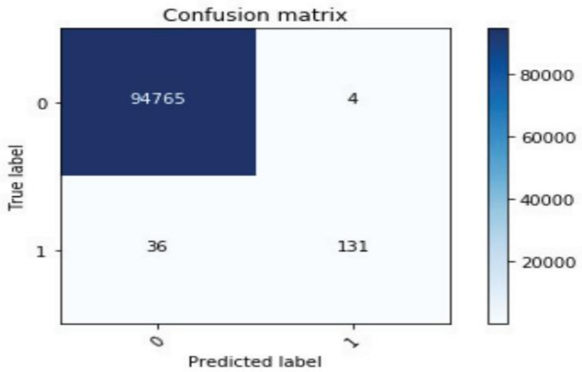

Figure 5: Confusion matrix for random forest algorithm

Figur 5 displays that the RF algorithm correctly guesses zeros in the end result 94765 times, but it gets them wrong 4 times. According to the RF method, it correctly guesses ones 131 times and wrongly 36 times.

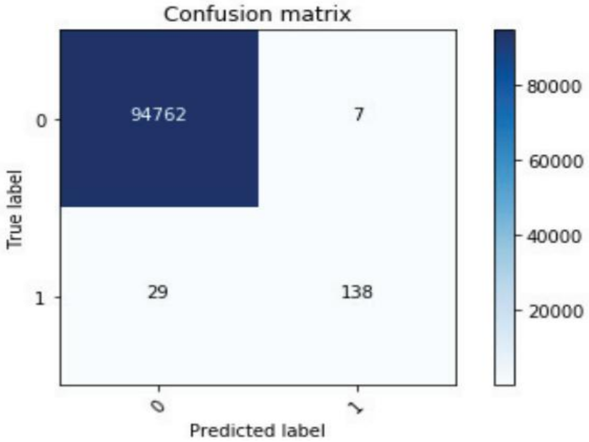

Figure 6: Confusion matrix for XGBOOST algorithm

Figure 6 displays that the XGBOOST method correctly guesses zeros in the end result 94762 times, but it gets them wrong 7 times. One's are accurately predicted by the XGBOOST algorithm 138 times, and wrongly predicted 29 times.

Table II: XGBoost and RF model performance for fraud detection

ML Algorithms	Random Forest	XGBOOST
Accuracy	99.95	99.96
Precision	99.92	99.12
Recall	99.82	80.68
F1-score	99.82	88.95

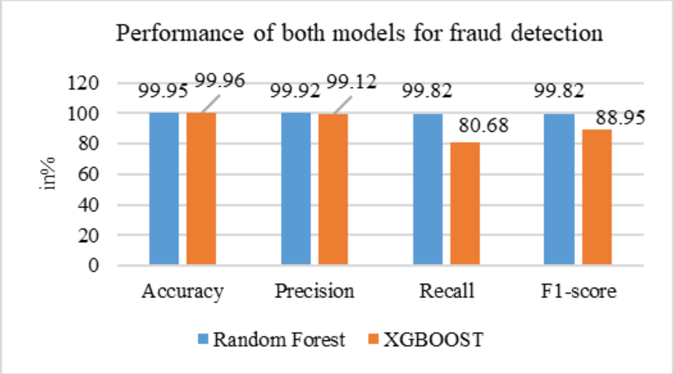

Figure 7: Parameters performance of RF and XGBoost models for fraud detection

In Table II and figure 7 shows, a performance of RF and XGBoost models for fraud detection is compared across key metrics. The RF model achieves an accuracy of 99.95%, closely followed by XGBoost at 99.96%. Precision is higher for RF at 99.92%, while XGBoost lags at 99.12%, indicating RF is better at minimizing false positives. In terms of recall, RF significantly outperforms XGBoost, with a score of 99.82% compared to 80.68%, showing RF's superiority in identifying actual fraud cases. The F1-score, a balance among precision and recall, also favors RF at 99.82%, compared to XGBoost's 88.95%, confirming RF's stronger overall performance for fraud detection.

Table III: Accuracy comparison of Machine Learning Algorithms for Fraud Detection

ML Algorithms	Accuracy (%)
Random Forest	99.95
XGBOOST	99.96
NN-[20]	90
Decision Tree -[21]	79.21

Table III shows how different ML models for finding banking fraud on credit card fraud data compare. In this compression, XGBoost leads the group with an impressive accuracy of 99.962%, closely followed by Random Forest at 99.957%. These algorithms demonstrate exceptional predictive capabilities, making them suitable for high-stakes applications. In contrast, Neural Networks (NN) show a significantly lower accuracy of 0.90%, indicating potential issues in model training or data quality. Additionally, the Decision Tree algorithm exhibits an accuracy of 79.21%, which, while better than NN, still falls short compared to the ensemble methods of Random Forest and XGBoost. Overall, the ensemble methods clearly outperform both NN and Decision Tree in this comparison, highlighting their effectiveness in achieving higher prediction accuracy.

V. CONCLUSION AND FUTURE SCOPE

The critical relevance of sophisticated technology in fighting financial crime is shown by comparing ML models with data analytics approaches for banking fraud detection. With a rise of sophisticated fraud schemes, financial institutions are turning to machine learning models like RF, DT, and XGBoost to identify suspicious activities in real time. This study demonstrates that

Random Forest and XGBoost outperform traditional methods, achieving impressive prediction accuracies of 99.957% and 99.962%, respectively. The application of techniques like anomaly detection, clustering, and predictive modeling has also proven effective in detecting fraudulent patterns within large datasets. These results highlight the potential of ML models to provide accurate and robust fraud detection solutions in banking, ensuring a secure financial environment for customers and institutions alike.

This framework has great potential for use in a huge range of interdisciplinary fields, like economic criminology, accounting, and law, while there are still many unanswered questions that need to be answered in the future. This study paves the way for further investigations on how to improve fraud detection systems. Researchers in these fields may find our study especially useful, as it provides a larger background that may inspire new lines of inquiry.

REFERENCES

1. Sadgali, N. Sael, and F. Benabbou, “Fraud detection in credit card transaction using machine learning techniques,” in ICSSD 2019 - International Conference on Smart Systems and Data Science, 2019. doi: 10.1109/ICSSD47982.2019.9002674.
2. S. Patil, V. Nemade, and P. K. Soni, “Predictive Modelling for Credit Card Fraud Detection Using Data Analytics,” in Procedia Computer Science, 2018. doi: 10.1016/j.procs.2018.05.199.
3. M. A. Ali, N. Hussin, and I. A. Abed, “E-banking fraud detection: A short review,” Int. J. Innov. Creat. Chang., vol. 6, no. 8, pp. 67–87, 2019.
4. K. Mullangi, N. D. Vamsi Krishna Yarlagadda, and M. Rodriguez, “Integrating AI and Reciprocal Symmetry in Financial Management: A Pathway to Enhanced Decision-Making,” Int. J. Reciprocal Symmetry Theor. Phys., vol. 5, no. 1, pp. 42–52, 2018.
5. Kulkarni, “Credit Card Fraud Detection Using Random Forest and Local Outlier Factor,” Int. J. Res. Appl. Sci. Eng. Technol., 2019, doi: 10.22214/ijraset.2019.4209.
6. J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, “Credit card fraud detection using machine learning techniques: A comparative analysis,” in Proceedings of the IEEE International Conference on Computing, Networking and Informatics, ICCNI 2017, 2017. doi: 10.1109/ICCNI.2017.8123782.
7. S. C. R. Vennapusa, T. Fadziso, K. Sachani, V. K. Yarlagadda, and S. K. R. Anumandla, “Cryptocurrency-Based Loyalty Programs for Enhanced Customer Engagement,” Technol. Manag. Rev., vol. 3, no. 1, pp. 46–62, 2018.
8. S. Venkata Suryanarayana, G. N. Balaji, and G. Venkateswara Rao, “Machine learning approaches for credit card fraud detection,” Int. J. Eng. Technol., 2018, doi: 10.14419/ijet.v7i2.9356.
9. S. Dhankhad, E. A. Mohammed, and B. Far, “Supervised machine learning algorithms for credit card fraudulent transaction detection: A comparative study,” in Proceedings - 2018 IEEE 19th International Conference on Information Reuse and Integration for Data Science, IRI 2018, 2018. doi: 10.1109/IRI.2018.00025.
10. N. Dornadula and S. Geetha, “Credit Card Fraud Detection using Machine Learning Algorithms,” in Procedia Computer Science, 2019. doi: 10.1016/j.procs.2020.01.057.
11. Jiang, J. Song, G. Liu, L. Zheng, and W. Luan, “Credit Card Fraud Detection: A Novel Approach Using Aggregation Strategy and Feedback Mechanism,” IEEE Internet Things J., 2018, doi: 10.1109/JIOT.2018.2816007.

12. K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit Card Fraud Detection Using AdaBoost and Majority Voting," IEEE Access, 2018, doi: 10.1109/ACCESS.2018.2806420.
13. Badejo, B. A. Okuneye, and M. R. Taiwo, "Fraud Detection in the Banking System in Nigeria: Challenges and Prospects," Shirkah J. Econ. Bus., vol. 2, 2018, doi: 10.22515/shirkah.v2i3.167.
14. O. S. Yee, S. Sagadevan, and N. H. A. H. Malim, "Credit card fraud detection using machine learning as data mining technique," J. Telecommun. Electron. Comput. Eng., vol. 10, no. 1–4, pp. 23–27, 2018.
15. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga, and N. Kuruwitaarachchi, "Real-time credit card fraud detection using machine learning," in Proceedings of the 9th International Conference On Cloud Computing, Data Science and Engineering, Confluence 2019, 2019. doi: 10.1109/CONFLUENCE.2019.8776942.
16. S. Josephine Isabella, S. Srinivasan, and G. Suseendran, "An Efficient Study of Fraud Detection System Using Ml Techniques," Lect. Notes Networks Syst., vol. 118, no. May, pp. 59–67, 2020, doi: 10.1007/978-981-15-3284-9_8.
17. dave and T. Adewale, "User-Centric Approaches to Fraud Detection Incorporating Behavioral Analytics in Azerbaijan's Banking Systems," 2021.
18. Jain, A. Purwar, and D. Yadav, "Credit Card Fraud Detection Using K-Means and Fuzzy C-Means," 2021. doi: 10.4018/978-1-7998-6870-5.ch016.
19. V. Jain, M. Agrawal, and A. Kumar, "Performance Analysis of Machine Learning Algorithms in Credit Cards Fraud Detection," in ICRITO 2020 - IEEE 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions), 2020. doi: 10.1109/ICRITO48877.2020.9197762.
20. K. Rai and R. K. Dwivedi, "Fraud Detection in Credit Card Data using Unsupervised Machine Learning Based Scheme," in Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020, 2020. doi: 10.1109/ICESC48915.2020.9155615.
21. S. Khatri, A. Arora, and A. P. Agrawal, "Card Fraud Detection : A Comparison," 2020 10th Int. Conf. Cloud Comput. Data Sci. Eng., pp. 680–683, 2020.